



UNITED STATES PATENT AND TRADEMARK OFFICE

UNITED STATES DEPARTMENT OF COMMERCE
United States Patent and Trademark Office
Address: COMMISSIONER FOR PATENTS
P.O. Box 1450
Alexandria, Virginia 22313-1450
www.uspto.gov

| APPLICATION NO. | FILING DATE | FIRST NAMED INVENTOR | ATTORNEY DOCKET NO. | CONFIRMATION NO. |
|-----------------|-------------|----------------------|---------------------|------------------|
| 09/615,726 | 07/13/2000 | Stephen R. Carter | 6647-13 | 4935 |

20575 7590 08/03/2004

MARGER JOHNSON & MCCOLLOM PC
1030 SW MORRISON STREET
PORTLAND, OR 97205

| |
|----------|
| EXAMINER |
|----------|

SPOONER, LAMONT M

| | |
|----------|--------------|
| ART UNIT | PAPER NUMBER |
|----------|--------------|

2654

DATE MAILED: 08/03/2004

8

Please find below and/or attached an Office communication concerning this application or proceeding.

Office Action Summary

Application No.

09/615,726

Applicant(s)

CARTER ET AL.

Examiner

Lamont M Spooner

Art Unit

2654

-- The MAILING DATE of this communication appears on the cover sheet with the correspondence address --

Period for Reply

A SHORTENED STATUTORY PERIOD FOR REPLY IS SET TO EXPIRE 3 MONTH(S) FROM THE MAILING DATE OF THIS COMMUNICATION.

- Extensions of time may be available under the provisions of 37 CFR 1.136(a). In no event, however, may a reply be timely filed after SIX (6) MONTHS from the mailing date of this communication.
- If the period for reply specified above is less than thirty (30) days, a reply within the statutory minimum of thirty (30) days will be considered timely.
- If NO period for reply is specified above, the maximum statutory period will apply and will expire SIX (6) MONTHS from the mailing date of this communication.
- Failure to reply within the set or extended period for reply will, by statute, cause the application to become ABANDONED (35 U.S.C. § 133). Any reply received by the Office later than three months after the mailing date of this communication, even if timely filed, may reduce any earned patent term adjustment. See 37 CFR 1.704(b).

Status

- 1) ☒ Responsive to communication(s) filed on 19 April 2004.
- 2a) ☐ This action is **FINAL**. 2b) ☒ This action is non-final.
- 3) ☐ Since this application is in condition for allowance except for formal matters, prosecution as to the merits is closed in accordance with the practice under *Ex parte Quayle*, 1935 C.D. 11, 453 O.G. 213.

Disposition of Claims

- 4) ☒ Claim(s) 1-30 is/are pending in the application.
- 4a) Of the above claim(s) _____ is/are withdrawn from consideration.
- 5) ☐ Claim(s) _____ is/are allowed.
- 6) ☒ Claim(s) 1-30 is/are rejected.
- 7) ☐ Claim(s) _____ is/are objected to.
- 8) ☐ Claim(s) _____ are subject to restriction and/or election requirement.

Application Papers

- 9) ☐ The specification is objected to by the Examiner.
- 10) ☒ The drawing(s) filed on 13 July 2000 is/are: a) ☒ accepted or b) ☐ objected to by the Examiner.
Applicant may not request that any objection to the drawing(s) be held in abeyance. See 37 CFR 1.85(a).
Replacement drawing sheet(s) including the correction is required if the drawing(s) is objected to. See 37 CFR 1.121(d).
- 11) ☐ The oath or declaration is objected to by the Examiner. Note the attached Office Action or form PTO-152.

Priority under 35 U.S.C. § 119

- 12) ☐ Acknowledgment is made of a claim for foreign priority under 35 U.S.C. § 119(a)-(d) or (f).
a) ☐ All b) ☐ Some * c) ☐ None of:
1. ☐ Certified copies of the priority documents have been received.
 2. ☐ Certified copies of the priority documents have been received in Application No. _____.
 3. ☐ Copies of the certified copies of the priority documents have been received in this National Stage application from the International Bureau (PCT Rule 17.2(a)).

* See the attached detailed Office action for a list of the certified copies not received.

Attachment(s)

- | | |
|--|---|
| 1) <input checked="" type="checkbox"/> Notice of References Cited (PTO-892) | 4) <input type="checkbox"/> Interview Summary (PTO-413) Paper No(s)/Mail Date. _____ |
| 2) <input type="checkbox"/> Notice of Draftsperson's Patent Drawing Review (PTO-948) | 5) <input type="checkbox"/> Notice of Informal Patent Application (PTO-152) |
| 3) <input checked="" type="checkbox"/> Information Disclosure Statement(s) (PTO-1449 or PTO/SB/08) Paper No(s)/Mail Date <u>7</u> . | 6) <input type="checkbox"/> Other: _____ |

DETAILED ACTION

Response to Arguments

1. Applicant's arguments with respect to claims 1-30 have been considered but are moot in view of the new ground(s) of rejection.

Claim Rejections - 35 USC § 102

2. The following is a quotation of the appropriate paragraphs of 35 U.S.C. 102 that form the basis for the rejections under this section made in this Office action:

A person shall be entitled to a patent unless –

(b) the invention was patented or described in a printed publication in this or a foreign country or in public use or on sale in this country, more than one year prior to the date of application for patent in the United States.

3. Claims 1, 3-6, 8-13, 15-17, 19, 21, 22, and 28 are rejected under 35 U.S.C. 102(b) as being anticipated by Caid et al. (US Patent No. 5,619,709 Apr. 8, 1997).

As per **claim 1**, Caid et al. discloses a method for determining dominant phrase vectors in a topological vector space (TVS) for a semantic content of a document on a computer system, the method comprising:

accessing dominant phrases for the document (C.5.lines 59-62-multiple word phrases, and single word phrases are included, C.6.lines 28-30), the dominant phrases representing a condensed content for the document (C.6.lines, 1-5);

constructing at least one state vector (C.6.lines 6-8) in the topological vector space (C.5.lines 27-29-the words in the geometrical space is interpreted as the TVS) for each dominant phrase using a dictionary and a basis (C.5.lines 16-19-the meaning of the word by dictionary, also C.27.lines 60-63, 65-67, and C.5.lines 28, 29-basis); and

collecting the state vectors into the dominant phrase vectors for the document (C.9.lines 9, 10-the summary vector includes the collected state vectors as dominant phrase vectors).

As per **claim 3**, Caid et al. discloses all of the limitations of claim 1, upon which claim 3 depends. Caid et al. further discloses:

accessing dominant phrases includes storing the dominant phrases in computer memory accessible by the computer system (C.9.lines 48, 49).

As per **claim 4**, Caid et al. discloses all of the limitations of claim 1, upon which claim 4 depends. Caid et al. further discloses:

forming a semantic abstract comprising the dominant phrase vectors (C.9.lines 7, 8).

As per **claims 5 and 10**, Caid et al. discloses a method for determining dominant vectors in a topological vector space or a semantic content of a document on a computer system, the method comprising:

storing the document in computer memory accessible by the computer system (C.4.lines 55-57).

extracting words for at least a portion of the document (C.6.lines 6-8-the extracted words are fed to the learning system);

constructing a state vector in the topological vector space (C.5.lines 27-29-the words in the geometrical space is interpreted as the TVS) for each word using a dictionary and a basis (C.5.lines 16-19-the meaning of the word by dictionary, also C.27.lines 65-67, and C.5.lines 28, 29-basis); and

constructing a state vector (C.6.line 36) in the topological vector space for each word using a dictionary and a basis (C.6.lines 46-50);
filtering the state vectors (C.8.lines 25-35); and
collecting the filtered state vectors into the dominant vectors for the document (Fig. 2 items 202, 203, C.6.lines 38-45).

As per **claim 6**, Caid et al. discloses all of the limitations of claim 5, upon which claim 6 depends. Caid et al. further discloses:

extracting words includes extracting words from the entire document (C.6.lines 6-8, 28-30).

As per **claim 8**, Caid et al. discloses all of the limitations of claim 5, upon which claim 8 depends. Carter et al further discloses:

filtering the state vectors includes calculating a centroid (C.9.lines 59-62-calculation step) in the TVS for the state vectors (C.9.lines 59-62-context vectors includes the state vectors); and

selecting the state vectors nearest the centroid (C.11.lines 32-44, C.12.lines 8, 9-the summary vector including the state vectors are selected and retrieved).

As per **claim 9**, Caid et al. discloses all of the limitations of claim 5, upon which claim 9 depends. Caid et al. further discloses:

forming a semantic abstract comprising the dominant vectors (C.9.lines 7,8-the summary vector represents the abstract, including the dominant vectors as included by the context vectors of all the stems).

As per **claims 11 and 16**, Caid et al. discloses a method for determining a semantic abstract in a topological vector space for a semantic content of a document on a computer system, the method comprising:

storing the document in a computer memory accessible by the computer system (C.4.lines 55-57);

determining dominant phrase vectors for the document (C.5.lines 50, 51-the terms represent the phrases, C.5.lines 60-62);

determining dominant vectors for the document (C.5.lines 50-51-the individual words represent the dominant vectors);

generating the semantic abstract using the dominant phrase vectors and the dominant vectors (C.9.lines 7, 8-the abstract is generated including dominant phrase vectors and dominant vectors).

As per **claim 12**, Caid et al. discloses all of the limitations of claim 11, upon which claim 12 depends. Caid et al. further discloses:

generating the semantic abstract includes reducing the dominant phrase vectors based on the dominant vectors (C.8.lines 22, 23-a phrase vector is separated, thereby reducing the phrase vector based on the dominant vector, C.9.lines 7, 8).

As per **claim 13**, Caid et al. discloses all of the limitations of claim 11, upon which claim 13 depends. Carter et al further discloses:

generating the semantic abstract includes reducing the dominant vectors based on the dominant phrase vectors (C.8.lines 30-32-the separate dominant vectors are reduced into a dominant phrase by converging the vectors, C.9.lines 7, 8).

As per **claim 15**, Carter et al discloses all of the limitations of claim 11, upon which claim 15 depends. Caid et al. further discloses:

identifying the lexemes or lexeme phrases (C.27.lines 60-63, 65-67-lexemes or lexemic phrases) corresponding to state vectors in the semantic abstract (C.12.lines 18-20, 21-23, 26-29).

As per **claims 17, 22 and 28**, Caid et al. discloses a method for comparing the semantic content of first and second documents on a computer system, the method comprising:

determining the semantic abstracts (C.9.lines 7, 8) for the first and second document (C.9.lines 24-26).

measuring a distance between the semantic abstracts (C.10.lines 50-52, 55, 56, 59-62, 65-67, C.11.lines 51-53-the distances between each summary vector/abstract about a centroid is computed, also C.28.lines 45-49, 56-48, 61, 62-mathematical measurement between the summary of multiple documents).

classifying how closely related the first and second documents are using the distance (C.28.lines 56-58-the closeness is classified hierarchically between documents).

As per **claim 19**, Caid et al. discloses all of the limitations of claim 17, upon which claim 19 depends. Caid et al. further discloses:

determining a centroid vector in the topological vector space for each semantic abstract (C.10.lines 50-52).

As per **claim 21**, Caid et al. discloses all of the limitations of claim 19, upon which claim 21 depends. Caid et al. further discloses:

measuring a Euclidean distance between centroid vectors (C.27.lines 60-63).

Claim Rejections - 35 USC § 103

4. The following is a quotation of 35 U.S.C. 103(a) which forms the basis for all obviousness rejections set forth in this Office action:

(a) A patent may not be obtained though the invention is not identically disclosed or described as set forth in section 102 of this title, if the differences between the subject matter sought to be patented and the prior art are such that the subject matter as a whole would have been obvious at the time the invention was made to a person having ordinary skill in the art to which said subject matter pertains. Patentability shall not be negated by the manner in which the invention was made.

5. Claims 2, 25-27, 29, and 30 are rejected under 35 U.S.C. 103(a) as being unpatentable over Caid et al. in view of Turney (U.S. Patent No. 6,470,307 filed Jun. 23, 1997).

Caid et al. and Turney are analogous art because they are both of the semantic textual information retrieval field.

As per **claim 2**, Caid et al. discloses all of the limitations of claim 1, upon which claim 2 depends. Caid et al. does not explicitly disclose:

accessing dominant phrases includes extracting the dominant phrases from the document using a phrase extractor.

However, Turney teaches extracting dominant phrases from a document using a phrase extractor (C.3.lines 50-55). Therefore, at the time of the invention, it would have been obvious to one ordinarily skilled in the art to combine Caid et al. with Turney. The motivation for doing so would have been to determine a short summary of the document comprising the extracted phrases (Turney C.3.lines 52-55).

As per **claim 25**, Caid et al. discloses an apparatus on a computer system to determine a semantic abstract in a topological vector space for a semantic content of a document stored on the computer system, the method comprising:

a state vector constructor adapted to construct at least one state vector in the topological vector space (C.5.lines 27-29-the words in the geometrical space is interpreted as the TVS, C.6.lines 6-8); and

collection means for collecting the state vectors into the semantic abstract for the document (C.9.lines 9, 10-the summary vector includes the collected state vectors as dominant phrase vectors).

Caid et al. does not explicitly disclose:

a phrase extractor adapted to extract phrases from the document

However, Turney teaches extracting dominant phrases from a document using a phrase extractor (C.3.lines 50-55). Therefore, at the time of the invention, it would have been obvious to one ordinarily skilled in the art to combine Caid et al. with Turney. The motivation for doing so would have been to determine a short summary of the document comprising the extracted phrases (Turney C.3.lines 52-55).

As per **claim 26**, Caid et al. and Turney disclose all of the limitations of claim 25, upon which claim 26 depends. Caid et al. further discloses:

filter means for filtering the state vectors to reduce the size of the semantic abstract (C.8.lines 17-19, 30-32-reducing the state vectors is interpreted as reducing the semantic abstract).

As per **claim 27**, Caid et al. and Turney disclose all of the limitations of claim 25, upon which claim 27 depends. Caid et al. further discloses:

the state vector constructor is further adapted to construct a state vector for each word in the document (C.6.lines 6-8, 28-30).

As per **claim 29**, Caid et al. discloses a method for determining a semantic abstract in a topological vector space for a semantic content of a document on a computer system, the method comprising:

the dominant phrases representing a condensed content for the document (C.6.lines 1-5);

constructing at least one first state vector (C.6.lines 6-8) in the topological vector space (C.5.lines 27-29-the words in the geometrical space is interpreted as the TVS) for each dominant phrase using a dictionary and a basis (C.5.lines 16-19-the meaning of the word by dictionary, also C.27.lines 60-63, 65-67, and C.5.lines 28, 29-basis, C.5.lines 50, 51-includes vector construction for phrases);

collecting the first state vectors into dominant phrase vectors for the document (C.5.lines 16, 17);

extracting words from at least a portion of the document (C.6.lines 6-8);

constructing at least one second state vector (C.6.lines 6-8) in the topological vector space (C.5.lines 27-29-the words in the geometrical space is interpreted as the TVS) for each word using a dictionary and a basis (C.5.lines 16-19-the meaning of the word by dictionary, also C.27.lines 60-63, 65-67, and C.5.lines 28, 29-basis, C.5.lines 50, 51-includes vector construction for each word);

filtering the second state vectors (C.8.lines 17-19, 30-32, 34-36);
collecting the filtered second state vectors into dominant vectors for the document (Fig. 2 items 202, 203, C.6.lines 38-45); and
generating the semantic abstract using the dominant phrase vectors and the dominant vectors (C.9.lines 5-10).

Caid et al. does not explicitly disclose:

extracting dominant phrases from the document using a phrase extractor
(C.4.lines 57, 58)

However, Turney teaches extracting dominant phrases from a document using a phrase extractor (C.3.lines 50-55). Therefore, at the time of the invention, it would have been obvious to one ordinarily skilled in the art to combine Caid et al. with Turney. The motivation for doing so would have been to determine a short summary of the documents comprising the extracted phrases (Turney C.3.lines 52-55), which are filtered and processed as taught by Caid et al.

As per **claim 30**, Caid et al. and Turney disclose all of the limitations of claim 29, upon which claim 30 depends. Caid et al. further discloses:

comparing the semantic abstract with a second semantic abstract for a second document (C.9.lines 7, 8, C.9.lines 24-26-second document) to determine how closely related the contents of the document are (C.28.lines 61, 62- the comparison thereof).

6. Claim 7 is rejected under 35 U.S.C. 103(a) as being unpatentable over Caid et al. in view of Kupiec et al. (U.S. Patent No. 5,778,397 Jul. 7, 1998).

Caid et al. and Kupiec are analogous art because they are both of the semantic textual information retrieval field.

As per **claim 7**, Caid et al. discloses all of the limitations of claim 5, upon which claim 7 depends. Caid et al. does not explicitly disclose:

filtering the state vectors includes selecting the state vectors that occur with highest frequencies.

However, Kupiec et al. teaches filtering keywords includes selecting the key words that occur with the highest frequencies (C.8.lines 8-10). Therefore, at the time of the invention, it would have been obvious to one ordinarily skilled in the art to combine Caid et al. with Kupiec et al. The motivation for doing so would have been to evaluate a direct theme, by filtering out less frequently used words, on the basis that frequently used words identify a documents theme (Kupiec C.4.lines 25-29).

7. Claims 20, 23, 24 are rejected under 35 U.S.C. 103(a) as being unpatentable over Caid et al. in view of Hazlehurst et al. (U.S. Patent No. 6,289,353 filed Jun. 10, 1999).

Caid et al. and Hazlehurst et al. are analogous art because they are both of the semantic textual information retrieval field.

As per **claim 20**, Caid et al. discloses all of the limitations of claim 19, upon which claim 20 depends.

Caid et al. does not explicitly disclose:

measuring a distance further includes measuring an angle between the centroid vectors.

However, Hazlehurst teaches measuring the distance between any two vectors can be done by the cosine function (C.12.lines 14-18). Therefore it would have been obvious to one ordinarily skilled in the art to combine Caid et al. with Hazlehurst et al. The motivation for doing so would have been to refine the distance measurement to a number using a very well known method of angle measurement in order to determine an exact measurement of distance between centroidal spaces.

As per **claim 23**, Caid et al. discloses a method for locating a second document on a computer with a semantic content similar to a first document, the method comprising:

- determining a semantic abstract for the first document (C.9.lines 7, 8);

- locating a second document (C.28.lines 49, 50-finding documents-the plural indicating a second document).

- determining a semantic abstract for the second document (C.9.lines 24-26);

- measuring a distance between the semantic abstracts for the first and second documents (C.10.lines 50-52, 55, 56, 59-62, 65-67, C.11.lines 51-53-the distances between each summary vector/abstract about a centroid is computed, also C.28.lines 45-49, 56-48, 61, 62-mathematical measurement between the summary of multiple documents).

- classifying how closely related the first and second documents are using the distance (C.28.lines 56-58-the closeness is classified hierarchically between documents).

- Caid et al. does not disclose;

Art Unit: 2654

if the second document is classified as having a semantic content similar to the semantic content of the first document, selecting the second document.

However, as it is well known in the art, Hazlehurst et al. teaches selecting the semantically related document (C.13.lines 57-61). Therefore it would have been obvious to one ordinarily skilled in the art to combine Caid et al. with Hazlehurst et al. The motivation would have been to select a document that was semantically related to another document by content, which would enable relevant retrieval or selection.

As per **claim 24**, Caid et al. and Hazlehurst et al. disclose all of the limitations of claim 23, upon which claim 24 depends.

Caid et al. does not disclose:

if the document is classified as not having a semantic content similar to the content of the first document, rejecting the second document.

However, as it is well known in the art, Hazlehurst et al. teaches of acquiring documents entailing specific conceptual content and discarding unpopular content (C.14.lines 8-12). Therefore it would have been obvious to combine Caid et al. with Hazlehurst et al. The motivation for doing so would have been to discard the documents that didn't relate well to the semantic similarities necessary for desired selection.

8. Claim 14 is rejected under 35 U.S.C. 103(a) as being unpatentable over Caid et al. in view of Arai et al. (U.S. Patent No. 6,173,261 filed Dec. 21, 1998).

Caid et al. and Arai et al. are analogous art because they are both of the semantic textual information retrieval field.

As per **claim 14**, Caid et al. discloses all of the limitations of claim 11, upon which claim 14 depends.

Caid et al. does not disclose:

obtaining a probability distribution function for a reduced set of the dominant phrase vectors similar to a probability distribution function for the dominant phrase vectors.

However, as it is well known in the art, Arai et al. teaches having a probability distribution function for phrases (C.2.lines 24-30). Therefore it would have been obvious to one ordinarily skilled in the art to combine Caid et al. with Arai et al. The motivation for doing so would have been to measure the similarity between the phrase vectors as it is possible to do so with the determination of the probability distributions of the phrases.

9. Claim 18 is rejected under 35 U.S.C. 103(a) as being unpatentable over Caid et al. in view Huttenlocher et al. (U.S. Patent No. 5,539,841 Jul. 23, 1996).

Caid et al. and Huttenlocher et al. are analogous art in that they are both involve semantic comparison of information.

As per **claim 18**, Caid et al. discloses all of the limitations of claim 17, upon which claim 18 depends.

Caid et al. does not disclose:

measuring a Hausdorff distance between the semantic abstracts.

However, Huttenlocher et al. teaches employing the Hausdorff distance as a measuring technique to determine the closeness of word images or semantic units/abstract of understanding (Abstract, C.12.lines 1-3). Therefore it would have been

Art Unit: 2654

obvious to one of ordinary skill in the art to combine Caid et al. with Huttenlocher et al.

The motivation for doing so would have been to determine how close two abstracts were in accordance with one another in a set space, because using the Hausdorff distance specializes in determining a measure for comparing point sets (Huttenlocher C.13lines 53-55).

Conclusion

10. The prior art made of record and not relied upon is considered pertinent to applicant's disclosure.

- Nakao (US Patent No. 6,205,456 filed Jan. 13, 1998) teaches summarizing a document, having an extraction unit for extracting dominant words and phrases.
- Liddy et al. (US Patent No. 5,873,056 Feb. 19, 1999) teaches having a semantic vector representation of a text.
- Sotomayor (US Patent 5,708,825 Jan. 13, 1998) teaches locating documents with similar semantic abstracts.
- Foltz et al. (US Patent No. 6,356,864 filed Jul. 23, 1998) teaches comparing vector spaces of documents.
- Vogel (US Patent No. 5,963,965 Oct. 9, 1999) teaches having a centroid and documents of relevance surrounding the centroid, for location of corresponding documents of relevant interest.
- Witbrock et al. (US Patent No. 6,317,708 filed Jul. 12, 1999) teaches generating semantic abstracts from documents.

Art Unit: 2654

11. Any inquiry concerning this communication or earlier communications from the examiner should be directed to Lamont M Spooner whose telephone number is 703/305-8661. The examiner can normally be reached on 8:00 AM - 5:00 PM.

If attempts to reach the examiner by telephone are unsuccessful, the examiner's supervisor, Nguyen Vo can be reached on 703/308-6728. The fax phone number for the organization where this application or proceeding is assigned is 703-872-9306.

Information regarding the status of an application may be obtained from the Patent Application Information Retrieval (PAIR) system. Status information for published applications may be obtained from either Private PAIR or Public PAIR. Status information for unpublished applications is available through Private PAIR only. For more information about the PAIR system, see <http://pair-direct.uspto.gov>. Should you have questions on access to the Private PAIR system, contact the Electronic Business Center (EBC) at 866-217-9197 (toll-free).

lms
07-16-04

Nguyen Vo
7-26-2004

NGUYENT.VO
PRIMARY EXAMINER